First publ. in: Information Visualization 2 (2003), 1, pp. 40-50

# Visual unrolling of network evolution and the analysis of dynamic discourse†

Ulrik Brandes[1],
Steven R. Corman[2]

[1]Department of Mathematics & Computer Science, University of Passau, 94030 Passau, Germany; [2]Hugh Downs School of Human Communication, College of Public Programs, Arizona State University, Tempe, AZ 85287-1205, U.S.A.

**Correspondence:**
**Ulrik Brandes, Department of Mathematics & Computer Science, University of Passau, 94030 Passau, Germany.**
**E-mail: brandes@algo.fmi.uni-passau.de**

†Research supported in part by a gift from Intel Corporation and by Deutsche Forschungsgemeinschaft (DFG) under Grant BR 2158/1-1.

**Abstract**
We introduce a method for visualizing evolving networks. In addition to the intermediate states of the network, it conveys the nature of change between states by unrolling the dynamics of the network. Each modification is shown in a separate layer of a three-dimensional representation, where the stack of layers corresponds to a time line of the evolution. We focus on networks of dynamic discourse as the driving application, but the method extends to any type of networks evolving in similar ways.

**Keywords:** Dynamic network visualization; graph drawing; network text analysis; centering resonance analysis; betweenness centrality

## Introduction

Evolving networks are networks that change over time, thus presenting a unique challenge to network visualization. They occur in diverse areas such as Web graph analysis, social network analysis, software engineering, and so on, and are often visualized using methods from dynamic graph drawing.[1,2] These methods typically generate a sequence of interdependent visualizations that represent intermediate states of the network, and try to preserve the user's mental map.[3] Graph animation (see, e.g., Friedrich and Houle[4]) is often used to lower the cognitive effort required to follow the transition from one visualization to the next.

However, if the actual changes occurring between consecutive states of an evolving network are an essential aspect of the data to be analyzed, it is not sufficient to compute interdependent layouts and animate between them, since animations nicely represent the transition between states, but do not allow a detailed scrutiny of change across several consecutive states. Animating an evolving network is similar to showing a time series as a point moving up and down rather than a graph. To facilitate simultaneous analysis of state and preceding changes, it is necessary to somehow unroll the process. Therefore, we propose to represent each step of a network's evolution as a layer in a three-dimensional visualization.

The problem is especially relevant to the domain of dynamic conversation and discourse. Researchers in a variety of social science disciplines are interested in the dynamics of conversations and other streams of discourse. By studying these dynamics they can better learn how meaning is created and shaped through interaction. Here, looking only at transitions between periodic states is like coming in on the middle of a conversation: to understand what is really going on, the analyst needs a visual representation that conveys a sense of the history of the process, similar to what a participant might have had.

The visualization approach introduced here has been developed with the analysis of dynamic discourse in mind, but it is straightforwardly applied

Konstanzer Online-Publikations-System (KOPS)
URN: http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-71847
URL: http://kops.ub.uni-konstanz.de/volltexte/2009/7184/

to any type of network that evolves similar ways. A recent example influenced by our ideas are visualizations of financial data.[5]

This paper is organized as follows. We begin in the next section by describing an existing method for network text analysis and its extension to dynamic discourse networks. The proposed approach to visualizing them and some illustrative examples are given in the section on Visually unrolling network evolution. In the penultimate section, we demonstrate the usage of our method on real-world data. Some conclusions are offered in the final section.

## Networks of dynamic discourse

In this section, we describe a recently introduced method for static text analysis and its extension to dynamic discourse such as human conversations, e-mail streams, or newsgroup archives. The method is described in some detail to emphasize the need for a visualization method that displays the nature of change that leads to new states of an evolving network.

### Static text

*Centering resonance analysis* (CRA)[6] is a method of network text analysis[7] that is designed for the study of complex discourse systems. This encompasses a wide range of phenomena, including interpersonal conversations, group discussions, interaction in large organizations, the Internet, and other mass media, as well as even larger social groups.

Scholars interested in the analysis of such discourse span a number of disciplines, including communication,[8] linguistics,[9] media studies,[10] organizational studies,[11] social psychology,[12] and sociology.[13] This is a diverse group that defies overgeneralization; however, it seems safe to say two things. First, these scholars all focus on how language is used to construct meaning within and across various episodes of communication. Discourse is inherently processual, so there is always a critical temporal element to these approaches. Second, the area is largely dominated by challenging qualitative research methods that call for either manual tracking of complex structures across multiple texts, or tedious transcription and analysis of conversation details and timing, including pauses, breaths, tone of voice and so on (see Schiffirin[14]). As a result, these analyses are almost always focused on very small samples of discourse.

CRA, and particularly its dynamic extension described below, offers an alternative to these laborious manual methods. It uses computational linguistics and linking rules to transform text (or transcribed conversation) into a network representation. The networks do not rely on context-specific semantic rules, training sets, or predefined document collections to analyze texts, but rather the method seeks to represent the author or speaker's communicative intent.[6] CRA therefore automatically produces stand-alone, abstract network representations of texts that can be analyzed alone or arbitrarily combined or compared with other CRA networks.

Let $T$ be a grammatically correct text. CRA constructs a graph $G(T) = (V,E)$ from $T$ in the following way:

1. The text is split into individual sentences.
2. For each sentence, noun phrases are identified using linguistic analysis.
3. A vertex is introduced for every noun[a] or adjective in a noun phrase.
4. An edge is introduced between every pair of vertices corresponding to words that occur in the same noun phrase, or are consecutive in the same sentence.

Put another way, small networks are constructed for each sentence, where words are considered linked if they co-occur inside noun phrases or occur on adjacent ends of consecutive noun phrases within that sentence. These networks are merged over all the sentences in the text.

The method thus yields a network of words comprising the subjects and objects of the text and how these are related to one another, and hence a representation of the text's structure. The graph $G(T)$ is called the *CRA network* of text $T$. An example of this construction is shown in Figure 1.

### Structural importance of words

Competent writers and speakers deploy words strategically to create a sensible, coherent message. Therefore, the structural positions of words in a CRA network $G = (V,E)$ reflect their importance or influence in structuring the meanings of the text. While many measures of structural position are conceivable, CRA uses a particular measure that reflects the extent to which a word is involved in chains of association between other words in the network.

Define the *betweenness centrality*[16,17] a vertex $v \in V$ as

$$c_v = \sum_{s \neq v \neq t \in V} \frac{\sigma_G(s,t|v)}{\sigma_G(s,t)},$$

where $\sigma_G(s,t|v)$ is the number of shortest paths between $s$ and $t$ passing through $v$, and $\sigma_G(s,t)$ the total number of shortest $(s,t)$-paths. Thus, a word with high betweenness is influential because it is greatly involved in channeling flows of meaning in the network.

In the example in Figure 1 vertices are placed at a distance from the center that reflects their betweenness centrality. Circles indicate the levels of centrality, and vertices outside the outer circle (representing the smallest non-zero centrality score) have zero centrality.

In a network with $n$ vertices and $m$ edges, betweenness centrality can be computed in $\mathcal{O}(nm)$ time and $\mathcal{O}(n+m)$ space as follows.[18] The length and number of shortest paths from a vertex $s \in V$ to all other vertices $t \in V$ is determined by breadth-first search starting from $s$. Instead of summing $n^2$ fractions, which would yield an

[a]We here disregard pronouns because the identity of particular individuals is not especially relevant to the analysis. See Corman *et al.*[6] for more information on this choice and its implications.

"One of the most essential branches of English liberty is the freedom of one's house. A man's house is his castle; and whilst he is quiet, he is as well guarded as a prince in his castle."

(a)

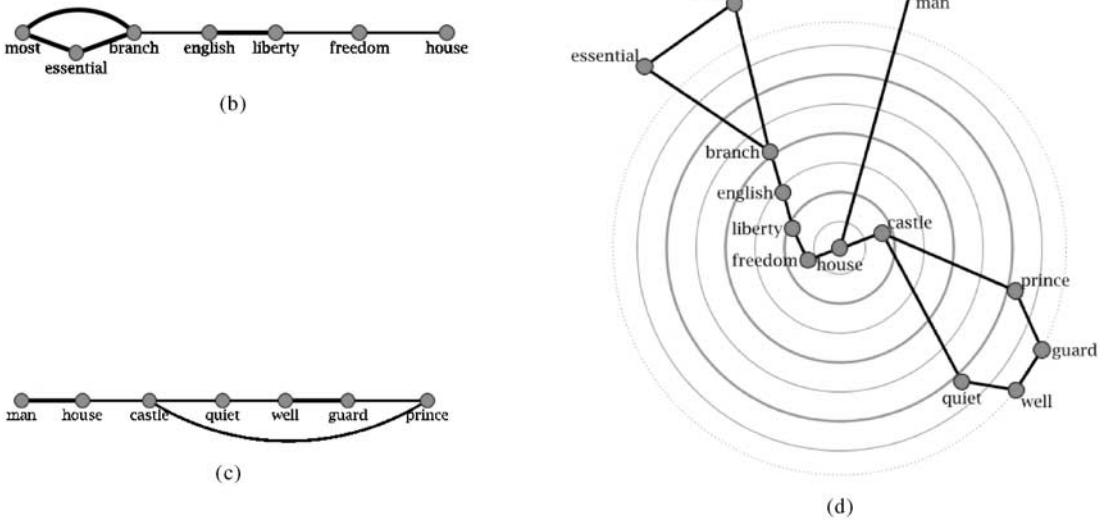



**Figure 1** CRA network extracted from a short text sample. The merged, static network is visualized with concentric circles that indicate levels of centrality.[15] Thus, 'house', 'freedom', and 'castle' are the most influential words. (a) sample text *T* with highlighted noun phrases (stemmed words in boldface) (b) first sentence (c) second sentence (d) CRA network *G*(*T*)

overall running time of $\mathcal{O}(n^3)$, we exploit recurrence

$$\delta_G(s|v) = \sum_{w:v\in P_s(w)} \frac{\sigma_G(s,v)}{\sigma_G(s,w)}(1+\delta_G(s|w)),$$

where $\delta_G(s|v) = \sum_{t\in V} \sigma_G(s,t|v)/\sigma_G(s,t)$ is the *dependency* of $s$ and $v$ and $P_s(w)$ is the set of all immediate predecessors of $w$ on shortest paths from $s$ to $w$. Since $c_v = \sum_{s\in V} \delta_G(s|v)$, betweenness can be accumulated from $n$ breadth-first searches, resulting in an overall running time of $\mathcal{O}(nm)$ rather than $\mathcal{O}(n^3)$. Moreover, space complexity is reduced from $\mathcal{O}(n^2)$ to $\mathcal{O}(n+m)$. See Brandes[18] for details and note that CRA networks tend to be sparse.

## Dynamic discourse

CRA has so far been applied to the analysis of static texts. However, discourse is best thought of as dynamic, involving a sequence of speaking turns or a set of time-ordered written texts (e.g., e-mails, press releases). One could analyze such a sequence of smaller texts by aggregating them into one large text and constructing a CRA network for it. However, change is a primary feature of interest, so analyzing a CRA network of the entire discourse neglects what happens within turns or texts and the relation of words between them and across sets of them. The basic CRA framework is therefore extended to *dynamic centering resonance analysis* (DCRA) as follows.[6]

A *discourse* is a sequence $D=(S_1, \ldots, S_k)$ of *epsiodes* $S_t$ that represent, for example, a single speaking turn or a serial item of written text. The straightforward generalization of analyzing the CRA network $G(S_t)$ of each episode ignores interdependences among them. It is a crucial feature of discourse that humans have (limited) *conversational memory*.[19] In a longer conversation things said recently are more relevant than those said long ago, and beyond a certain point in the conversation the utterances are irrelevant (or nearly so) to what is being said at present. Thus, we use a parameter $d$ to represent the depth of this conversational memory, or more generally dependence, appropriate in a given context.

A parameter $d$ is therefore used to represent the *depth* of conversational memory, or more generally dependence, appropriate in a given context. At any given point in time, say $t$, $1\leqslant t\leqslant k$, we define the *state* of the discourse to be $T_t=(S_{t-d'+1},\ldots,S_t)$, where $d'=\min\{d, t\}$, that is the text obtained from concatenating episodes in the conversational memory. The importance of words is relative to

conversational memory, and hence calculated in each $G(T_t)$. Thus, importance is dynamic and relative to the changing circumstances of the conversation. The sequence $G(D) = (G(T_1), \ldots, G(T_k))$ constitutes the evolving *discourse network*.

## Visually unrolling network evolution

Evolving networks are characterized by the (in general gradual) change of the state of a network between consecutive points in time. Standard methods for dynamic graph visualization generate a drawing for each such state (taking into account the drawings of previous states) and animate between consecutive states, thus maintaining the user's mental map and smoothing the transition from one state to another.

However, even the most recent history of a state, and thus the nature of change that occurred, cannot be recognized from that state's visualization. We therefore propose a visualization method that allows us to examine simultaneously the state of a network and the changes that led to it by unrolling the sequence of events.

### Evolutionary cross-sections

Networks evolving in discrete steps are composed of subnetworks that exist at particular points in time. Since the abstract information inherent to these networks consists of both the structure of subnetworks and changes between them, we propose a layered visualization design in which each newly introduced part of the network is shown in a layer of its own, and the composition of consecutive layers represents the corresponding state of the network. In other words, we unroll the dynamics of networks into its constituent layers.

In dynamic discourse, each episode induces a modification of the current state through the addition of edges, the potential creation of vertices, and the deletion of edges and vertices that drop out of conversational memory. The layer introduced into our visualization at time $t$ therefore shows only the elements induced by the corresponding episode, that is $G(S_t)$. By displaying layers exactly for those episodes that make up the current conversational memory, and by making these layers semitransparent, we obtain a visual representation that can be viewed from the top to infer the entire state $G(T_t)$ of the conversation. However, it also shows the individual contribution of each epsiode $S_{t-d+1}, \ldots, S_t$ with past epsiodes slowly fading away. In Figure 2, each sentence of the quote in Figure 1 is interpreted as an epsiode of its own.

### Layout

Any graph layout algorithm might be used to determine the layout of layers and, in fact, the algorithm should be chosen in close accordance with the substantive background of the application generating the graph.

In principle, there is no look-ahead in a conversation and the layout of each layer should be contingent only on the layout of preceding layers, that is an algorithm for on-line dynamic graph layout is sought.[20] However, our column-like representation of vertices prevents subsequent adjustments of positions, so that only newly introduced vertices can be placed freely, while all others are fixed at their position in previous layers.

Since this highly constrained, serial layout results in poor readability of subsequent layers, we take a different approach. At least today, dynamic discourse analysis is not carried out in real time, so that all slices of a discourse are available from the beginning. We can therefore compute a layout of the CRA network $G(D)$ of the entire discourse $D = (S_1, \ldots, S_k)$, and use its vertex positions for the layout of layers $G(S_t)$, $1 \leqslant t \leqslant k$.

The global static layout algorithm chosen is a spring embedder variant that corresponds to metric multidimensional scaling of pairwise distances.[21] Such layouts graphically support the intuition of words being close or distant in terms of the structure of the discourse. See Figure 2(b) for an example. The algorithm's objective function is to minimize the squared differences between distances in the layout and distances in the graph, weighted by the square of that distance. This is known as local scaling, because it reduces the influence of long-range relations. For $G = (V,E)$ and vertex positions $(x_v)_{v \in V}$ we want to minimize

$$\sum_{u,v \in V} \frac{(||x_u - x_v|| - d_G(u,v))^2}{d_G(u,v)^2},$$

where $d_G(u,v)$ is the length of a shortest path between $u$ and $v$ in $G$. This is an $\mathcal{NP}$-hard task, so that gradient methods moving one vertex at a time are used to obtain a local minimum. A straightforward implementation runs in time $\mathcal{O}(n^2)$ per iteration, but typically few iterations are required.

Rather than starting with a random layout of the entire graph, it is useful to compute a preliminary layout of a prefix of the evolving sequence and add one slice at a time, with a few iterations of the gradient method to obtain a layout for the prolonged prefix. The benefit is two-fold. First, it has been shown experimentally that incremental addition of vertices improves the overall performance of the layout algorithm.[22] And second, it takes into account the evolution of the graph by increasing the influence of early states without neglecting the later states. This modification thus combines the advantages of global and incremental layout according to the evolution of the graph. We are now experimenting with other types of layout algorithm to further improve the substantive interpretability of layouts, and the efficiency of their computation.

### Additional attributes

Several aspects of a layered visualization can be varied to enrich its information content. In the context of discourse networks, the most interesting additional information are the speakers in each turn, and the changing importance of words.

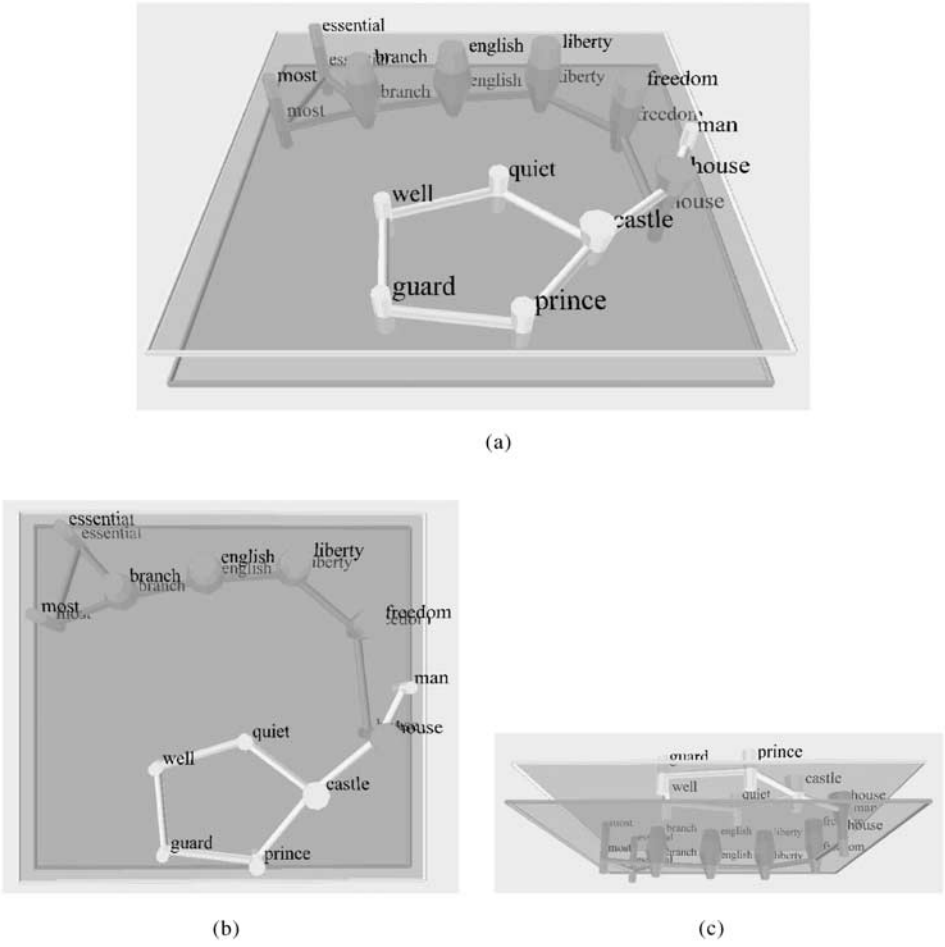



**Figure 2** Quote from Figure 1 treated as if sentences were from different speakers. The contents are almost disjoint, but linked through 'house'. Hence, 'house' becomes more influential after the second sentence, whereas most other words show decreasing importance. (a) stacked episodes of dynamic discourse (b) topic separation, fading in conversational memory (c) changing importance of words

We therefore assign a distinct color to each speaker and border each layer according to whose turn it represents. An edge is colored according to the speaker inducing it. Finally, the color of a speaker that introduces a word into the conversational memory is used for the corresponding vertex. This is illustrated in Figure 2(a).

As outlined in Section 2.2, the current importance of a word is captured by its betweenness centrality in the current state of the conversation. The change in centrality is conveyed by varying the thickness of columns that represent vertices. Let $c_-$ and $c_+$ be the minimum and maximum centrality score over the entire discourse, then the radius of a vertex $v$ with centrality $c_v(t)$ at time $t$ is defined as

$$r_- + \frac{c_v(t) - c_-}{c_+ - c_-}(r_+ - r_-),$$

where $r_-$ and $r_+$ denote the minimum and maximum desired thickness of vertices, so that the entire interval is actually used. In a top-down view, this results in central words being larger than peripheral words. Figure 2(c) shows how, when viewed from the side, the changing thickness of a column depicts the centrality profile of a word.

To reduce the complexity of large data sets, an importance threshold can be specified to exclude less important words from the visualization and emphasize the main structure of the discourse.

## Application examples

To evaluate our approach for exploratory analysis of evolving networks, a visualization system using Java and yFiles,[23] a library for drawing graphs, is implemented. The current prototype produces interactive visualizations through the extended application interface (EAI) for VRML, so that users can move forward and backward in the discourse and rotate, zoom, and pan the scene to view it from any point and orientation. In addition, each state is shown in a two-dimensional visualization that can be manipulated by the user, for example to highlight certain aspects by changing the layout, switch individual vertex

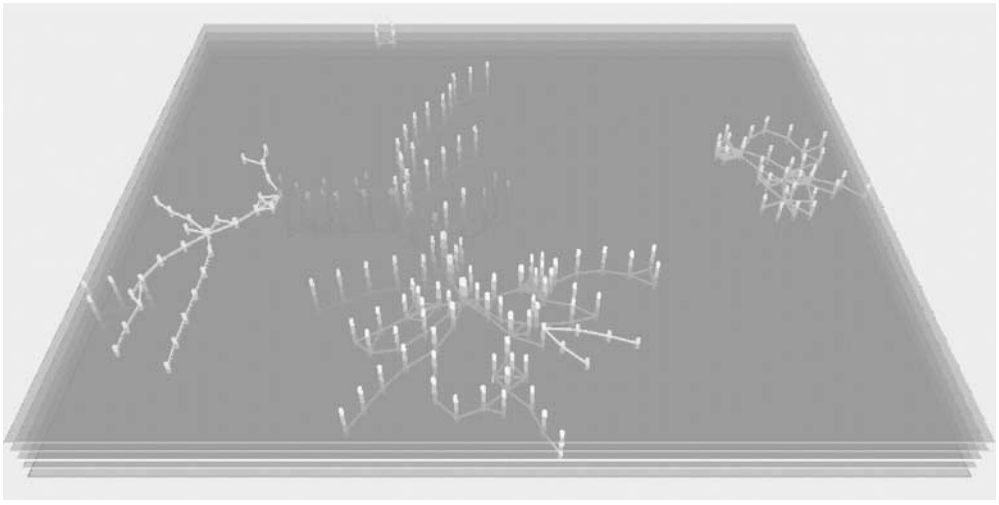


**Figure 3** Incoherent discourse.

labels on and off, or archive the current state in a VRML file.

The following are examples in which the system is used to support DCRA on recorded conversational data. Any other type of incrementally evolving networks, in which states are defined by cross-sections, would be appropriate as well. While the first example is crafted, the other two are extracts of real discourses.

### Incoherence

An important feature of discourse is coherence. If we have a sequence of random statements that make sense in themselves but have nothing to do with one another, these cannot form a coherent body of discourse. Visual DCRA should be able to identify these cases, and not create a lot of artifacts that make them appear integrated or sensible. To test for this we created a set of five arbitrary clips from unrelated sources and treated these as episodes of a discourse. The results are shown in Figure 3. These clips clearly occupy distinct areas in the layout and colors representing them are mostly confined to their own episode. There are only a few cases where layers connect due to chance common occurrence of the same word.

### Conversation

The first recorded example demonstrates the ability of our visualizations to indicate significant interventions by participants in a conversation. We analyzed a transcript of a video-taped conversation between Osama bin Laden and a man identified as 'Shaykh', which was released in December 2001.[b] The conversation begins with pleasantries and is structured almost exclusively by 'Shaykh'. After several turns, bin Laden makes an abstract reference to Koranic stories, thus marking a topic shift to reminiscing about the World Trade Center attacks. The DCRA visualization in Figure 4 shows this transition clearly. The dark nodes introduced by 'Shaykh' can still be seen on recent layers, but the top layer has words introduced by bin Laden, shown in light color, that connect to only two of 'Shaykh''s words ('allah' and 'people'). Otherwise, new words are introduced and links are mostly between those.

The examples in Figures 5 and 6 display other crucial excerpts from the conversation. While the first is probably the most frequently cited statement in the conversation, the second is an example of how a theme is picked up and reflected upon from a different angle.

### Dispute

The speakers in the second application are not exchanging information, but try to keep the upper hand in a dispute before an audience. Hence, the nature of the discourse should be different, too.

The following example consists of eight turns from the US Supreme Court deliberations in the Bush *vs* Gore case that decided the 2000 elections (Table 1). They constitute an exchange between Justice Kennedy (light) and Vice President Gore's lawyer, Mr. Boies (dark), at the very beginning of Boies's arguments. The original text of the turns is given in Figure 1, and Figure 7 shows two different views of turns 2–8.

Figure 7(a) shows the two-dimensional view of the entire conversational memory. Its layout has been modified manually to eliminate edge crossings and to improve the spatial grouping of words used by the two participants. The most influential words in these turns are 'florida' and 'supremecourt'. Also influential are 'cognizance' (linked to 'supremecourt') as well as 'law' and 'opinion' (both of which are linked to 'florida'). So at

[b]The transcript has been published by CNN at http://www.cnn.com/2001/US/12/13/tape.transcript. The first segment of the transcript was subjected to DCRA, with small disconnected components of the network (paths of length at most four) having been removed and evaluated separately.

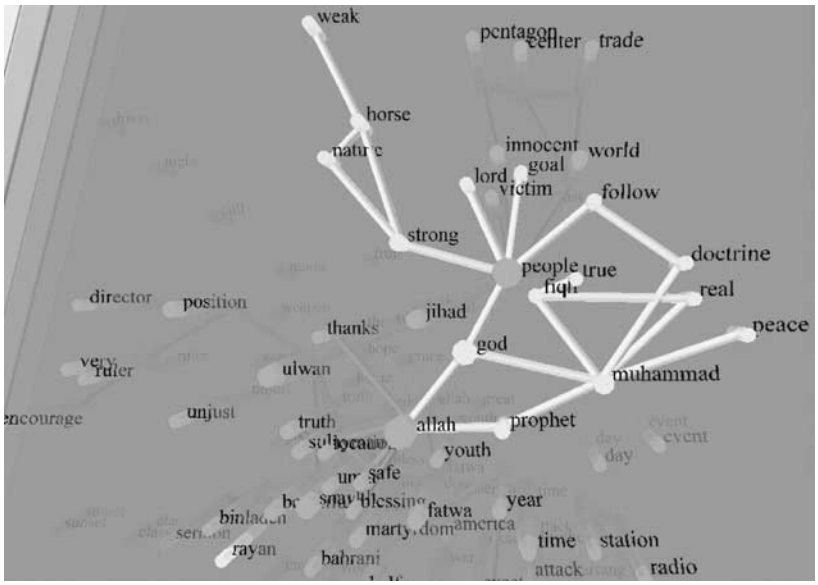



**Figure 4** bin Laden initiating a topic shift.

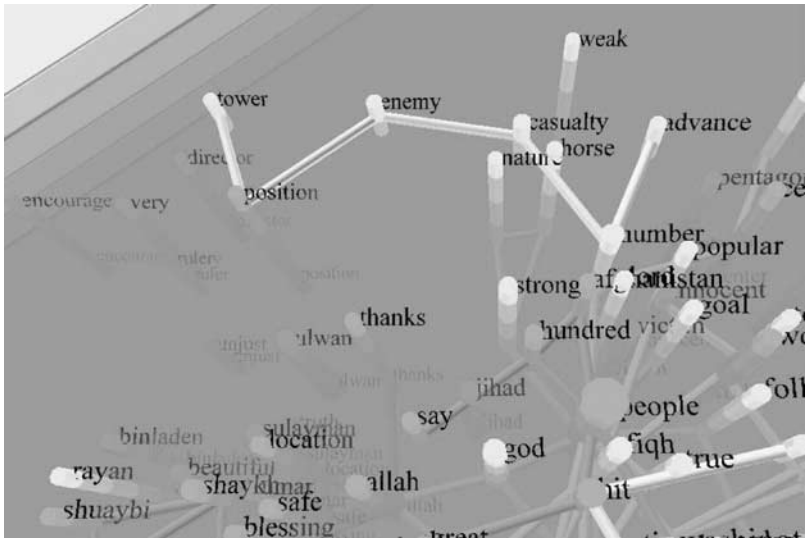



**Figure 5** '…we calculated in advance the number of casualties from the enemy…'.

a gross level we can identify the discourse as being about Florida law, the Florida opinion, and cognizance of its Supreme Court.

The evolution of this network is unrolled in Figure 7(b). It amply illustrates our point about the importance of the unrolling operation for this kind of analysis. Looking beneath the top layer, we can see that the words on the left side are all associated with Kennedy, and those on the right are all associated with Boies. The words associated with Boies are about the Florida opinion, its quality, and the cognizance of the Florida Supreme Court. The words associated with Kennedy, on the other hand, are about the law, especially changes in the law with respect to language in 'threeusc' (a US federal law related to the case), and a new law related to the deadline. Again excluding the top turn, the only nodes they use in common are: 'federal', 'deadline', 'florida', and 'supremecourt'. So it appears that Boies was making an effort to keep the discussion on his issues and not really engaging Kennedy's inquiries.

Finally, in the last (top) turn, see Boies making a statement that connects to 'language' on Kennedy's side. This same statement talks about the issue of a new law, finally connecting to Kennedy's issues. This giving in to Kennedy's queries marks a change point in the discussion, because on the next turn Chief Justice Rhenquist takes over with more technical questions about 'sectionfive' of 'threeusc'.

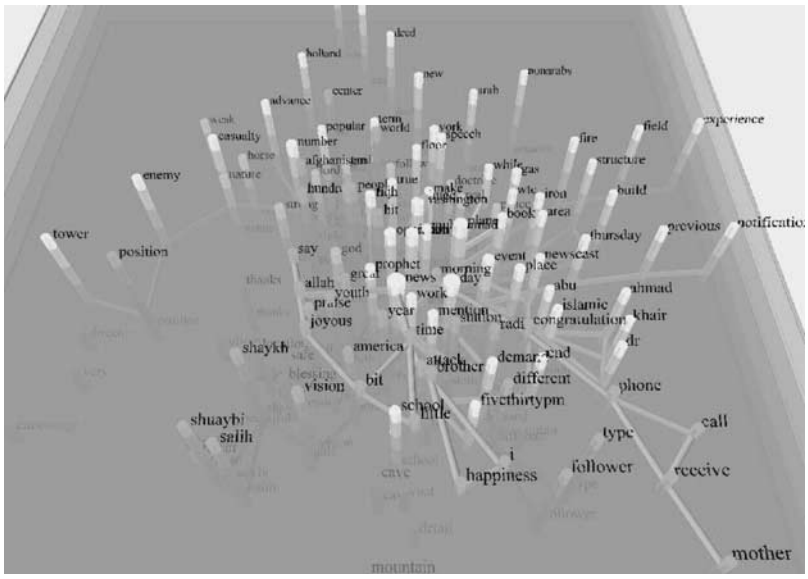


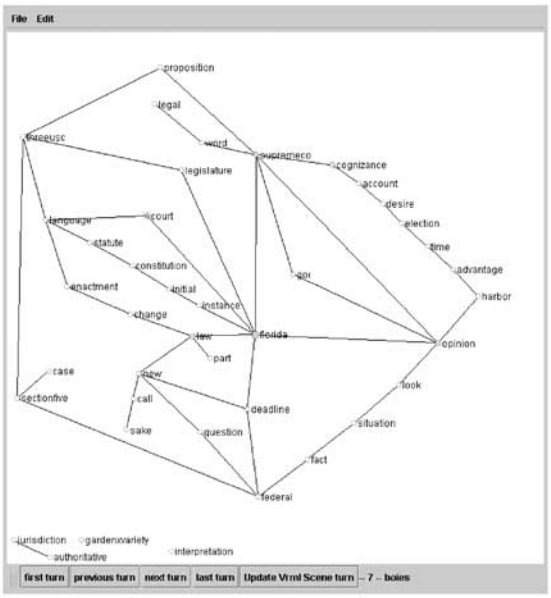

**Figure 6** After bin Laden has completed his account of the 'event', 'Shaykh' reflects on coverage in the 'news' and other reactions.

**Table 1 Eight episodes from the Bush vs Gore case**

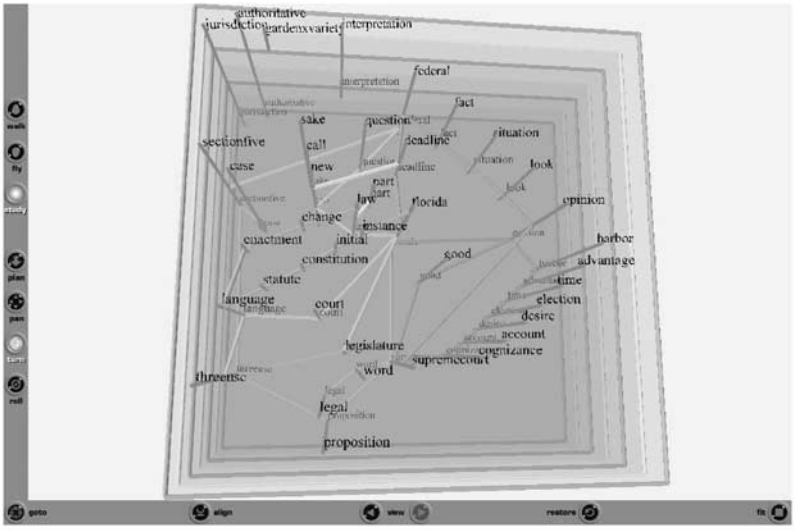
| *Speaker* | *Text* |
|---|---|
| BOIES | Thank you. Mr. Chief Justice, may it please the court. Let me begin by addressing what happened in the Beckstrom case that Mr. Klock refers to. |
| KENNEDY | Could we begin with jurisdiction first? The Supreme Court of Florida said that it was cognizant, and the legislature was cognizant, of 3USC Section 5. And for convenience sake, let's call that new law. That's not exactly this, but when the Supreme Court used that word, I assumed it used it in a legal sense. Cognizance means to take jurisdiction of, to take authoritative notice. Why doesn't that constitute an acceptance by the Supreme Court of the proposition that 3USC Section 5 must be interpreted in this case? |
| BOIES | I think, your honor, and, obviously, this court and the Florida Supreme Court is the best interpreter of that opinion, but I think a reasonable interpretation of that opinion is to say that what the Florida Supreme Court meant by cognizance is that it was taking into account the desire to get the election over in time so that everyone would have the advantage of the safe harbor, and I think that goes throughout the opinion. |
| KENNEDY | The language used in 3USC Section 5 is garden-variety language, so far as the courts are concerned. We can determine whether or not there is a new law or an old law. That's completely susceptible of judicial interpretation, is it not? [*Omitted: BOIES: Yes, I think it is, your honor.*] All right. And it seems to me that if the Florida court, and presumably the Florida legislature, have acted with reference to 3USC Section 5, that it presents now a federal question for us to determine whether or not there is or is not a new law by reason of the two Florida Supreme Court decisions? |
| BOIES | Except, your honor, what the Florida Supreme Court did, I think, in its opinion is to say that in terms of looking at how to remedy the situation, it needed to be cognizant of the fact that there was this federal deadline out there that was going to affect Florida's electors if that deadline was not met. |
| KENNEDY | Well, of course, the deadline is meaningless if there's a new law involved, and that's part of the equation, too. |
| BOIES | Yes, but what I would say is that whether or not there is a new law, that is, whether there is a change in the enactment in the language of the statute or the Constitution, is something that has to be decided in the initial instance by the Florida Supreme Court interpreting Florida law. |
| REHNQUIST | Mr. Boies, there are really two parts to that sentence of Section 5. One is the law in effect at the time, and the other is, finally determined six days before the date for choosing electors. Do you think the Florida Court meant to acknowledge? It seems to me since it's cited generally, they must have acknowledged both of those provisions. |

## Discussion

A different kind of change in discourse can be described more as gradual movement than wholesale change or avoidance. Here, we are interested in processes where an idea expressed at one point in a conversation is linked to an idea at a subsequent point in the conversation, and so on in a chaining-out fashion. In a DCRA visualization, such a structure appears as a staircase-like configuration. A brief real-life example of this form comes from a transcript of a graduate seminar discussion on group decisions support systems (GDSS). Figure 8 shows one part of the discussion where the issue was whether results




(a)




(b)

**Figure 7** Structured dispute in court. (a) conversational memory (b) conversational memory unrolled.

from GDSS studies using students are valid for real organizational settings.

In the layer below the surface, a student had made a rather complicated point about how students were unlikely to feel the same pressure to make quality decisions using the technology as in real organizations. On the top layer the instructor draws out the point that students are more willing to engage in quick processes to get their decisions made. This forms a small staircase, 'technology'–'student'–'quick'–'process', representing a movement in the conversation from students' relation with the technology to their relation with a particular kind of process.

## Conclusions

We have presented a visualization approach for evolving networks and applied it to visualize networks created by DCRA.

Based on the above examples, we conclude that DCRA visualizations are capable of indicating prominent features of discourse. They can show major interventions where a participant in a conversation changes the

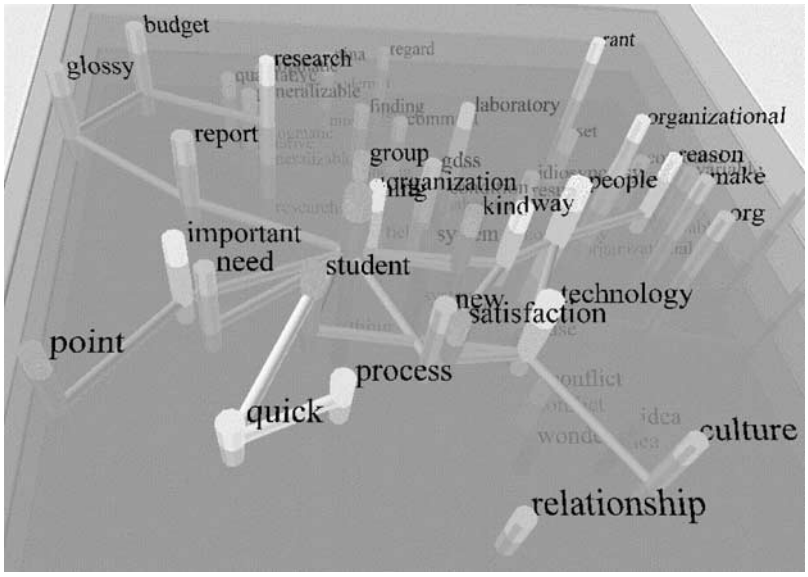



**Figure 8** Topical movement in group discussion illustrated by a 'staircase' structure: 'technology'–'student'–'quick'–'process'.

direction of a conversation. They can also show the more subtle movement involved in chaining ideas across turns. Finally, they can indicate cases where discourse is incoherent. It should be noted, though, that these conclusions have not yet been validated empirically.

Assessment of its potential for detecting more complex features of discourse awaits further research. For instance, it is planned to apply seriously visual DCRA to courtroom negotiations to identify or highlight strategies used by lawyers in particular cases. Potential other applications include newsgroup postings, organizational communication, and media streams. The sequence of importance scores for topics generated by DCRA can also be used as input for visualization approaches that do not depict network structure (see, e.g., Havre *et al*.[24]).

We see two implications for a fully developed DCRA visualization system. First, it could work as a tool in support of existing manual analysis methods. DCRA can tell us interesting things about the way a stream of discourse is unfolding, but it is not a substitute for interpretation by a human when complex questions of meaning are at stake. However, even in these situations it could be of considerable use to an analyst by identifying change points in the conversation, showing variation in the influence of different words, and so on.

Second, DCRA could enable new kinds of research on discourse. One of the chief limitations of existing methods is the degree of 'manual labor' involved in the analysis. As an automatic method, DCRA could allow quick analysis of dynamics in very broad systems of discourse. In addition, some of the patterns demonstrated in this paper could reveal new things about these systems. For example, the 'stairstep' linking of words across layers is a structure that represents the chaining of ideas across time and sources. Identifying these in the discourse of something like a scientific discipline should provide insight into the way new ideas develop.

One of the most interesting new applications would be the analysis of real-time conversation by coupling DCRA with voice recognition technology. We can envision a system, for instance, that monitors meetings and advises the group when its discussion is becoming incoherent, or has stayed on a single subject for too long. Unfortunately, although voice recognition technology has matured to the point that it makes a useful input device for some scientific applications under controlled conditions,[25] accurate speaker-independent continuous speech recognition remains in the future.[26] We hope to develop DCRA to exploit these advances as they arise.

Visual unrolling exploits a characteristic property of dynamic discourse and similar types of evolving networks, namely that change occurs only through the addition of new edges and vertices, and the deletion of the least recently introduced vertices and edges in the current state. It will be interesting to extend our method to the types of networks that evolve with arbitrary deletions.

Another direction for future research is the extension to real-time situations, for example by using dynamic graph layouts methods that produce suitable layouts of layers subject to many fixed positions inherited from earlier layers, or by relaxing the latter constraint and allowing vertex-representing columns to bend.

## Acknowledgments

We thank Odilo Oehmichen for implementing substantial parts of the DCRA visualization prototype.

## References


1 Herman I, Melançon G, Marshall MS. Graph visualization and navigation in information visualization: a survey. *IEEE Transactions on Visualization and Computer Graphics* 2000; **6**: 24–43.
2 Branke J. Dynamic graph drawing. In: Kaufmann M, Wagner D (Eds). *Drawing Graphs: Methods and Models*, Lecture Notes in Computer Science, Vol. 2025. Springer: Berlin, 2001; 228–246.
3 Misue K, Eades P, Lai W, Sugiyama K. Layout adjustment and the mental map. *Journal on Visual Languages and Computing* 1995; **6**: 183–210.
4 Friedrich C, Houle ME. *Graph drawing in motion II*. Ninth International Symposium on Graph Drawing (GD '01), Lecture Notes in Computer Science, Vol. 2265. Springer: Berlin, 2002; 220-231.
5 Dwyer T, Eades P. *Visualising a fund manager flow graph with columns and worms*. Sixth International Conference on Information Visualisation (IV '02), IEEE Computer Society Press: Silver spring, MD, 2002; 147–152.
6 Corman SR, Kuhn T, McPhee RD, Dooley KJ. Studying complex discursive systems: Centering resonance analysis of communication. *Human Communication Research* 2002; **28**: 157–206.
7 Carley KM. Network text analysis: the network position of concepts. In: Roberts C W. (Ed). *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts*. Lawrence Erlbaum Associates: London, 1997; 79–100.
8 Hopper R. *Telephone Conversation*. Indiana University Press: Bloomington, IN, 1992.
9 Johnstone B. *Discourse Analysis*. Blackwell: Oxford, 2002.
10 Bell A, Garett P. (Eds). *Approaches to Media Discourse.* Black-well: Oxford, 1998.
11 Boden D. *The Business of Talk: Organizations in Action*. Polity Press: Cambridge, MA, 1994.
12 Gergen KJ. *An Invitation to Social Construction.* Sage Publications: Beverley Hills, CA, 1999.
13 Silverman D. *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction.* Sage Publications: Beverley Hills, 2001.
14 Schiffrin D. *Approaches to Discourse*. Blackwell: Oxford, 1994.
15 Brandes U, Kenis P, Wagner D. Communicating centrality in policy network drawings. *IEEE Transactions on Visualization and Computer Graphics*, 2003; **9**(2): to appear.
16 Anthonisse JM. The rush in a directed graph. Technical Report BN 9/71, Stichting Mathematisch Centrum, Amsterdam, October. 1971.
17 Freeman LC. A set of measures of centrality based on betweenness. *Sociometry* 1977; **40**: 35–41.
18 Brandes U. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology* 2001; **25**: 163–177.
19 Stafford L, Waldron VR, Infield LL. Actor-observer differences in conversational memory. *Human Communication Research*. 1989; **15**: 590–611.
20 Brandes U, Wagner D. *A Bayesian paradigm for dynamic graph layout*. Fifth International Symposium on Graph Drawing (GD '97), Lecture Notes in Computer Science, Vol. 1353, Springer: Berlin, 1997; 236–247.
21 Kamada T, Kawai S. An algorithm for drawing general undirected graphs. *Information Processing Letters* 1989; **31**: 7–15.
22 Cohen JD. Drawing graphs to convey proximity: an incremental arrangement method. *ACM Transactions on Computer-human Interaction*. 1997; **4**: 197–229.
23 Wiese R, Eiglsperger M, Kaufmann M. *yFiles: visualization and automatic layout of graphs.* Ninth International Symposium on Graph Drawing (GD '01), Lecture Notes in Computer Science, Vol. 2265: Springer: Berlin, 2002; 453–454.
24 Havre S, Hetzler E, Whitney P, Nowell L. ThemeRiver: visualizing thematic changes in large document collections. *IEEE Transactions on Visualization and Computer Graphics* 2002; **8**: 9–20.
25 White DJ, King AP, Duncan SD. Voice recognition technology as a tool for behavioral research. *Behavior Research Methods, Instruments, & Computers* 2002; **34**: 1–5.
26 Brown EW, Srinivasan S, Coden A, Ponceleon D, Cooper JW. Toward speech as a knowledge resource. *IBM Systems Journal* 2001; **40**: 985–1001.